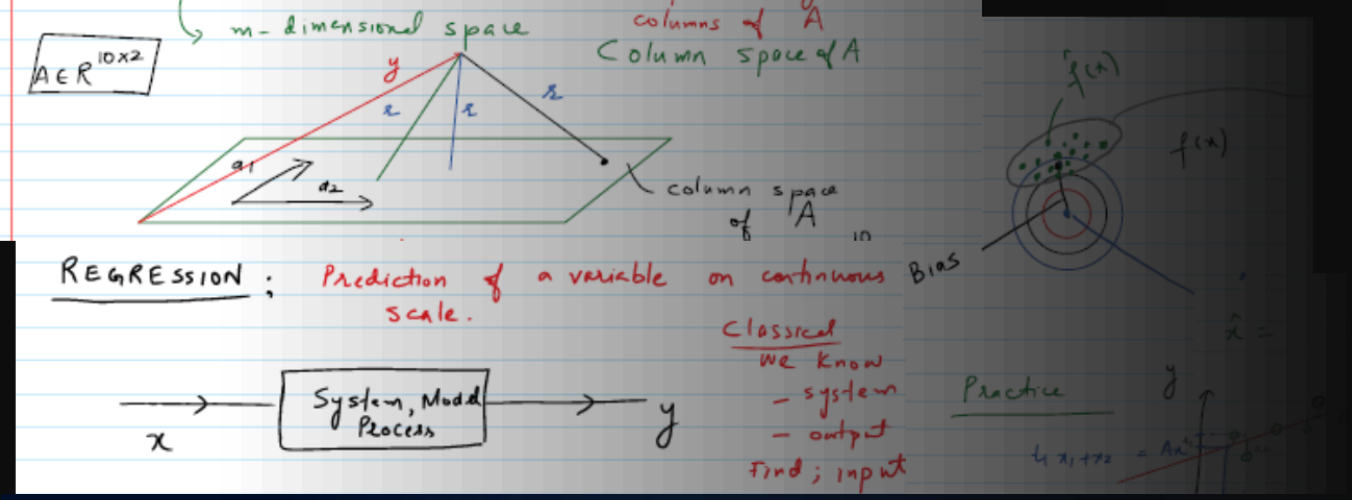


Machine Learning

EE514 – CS535

Overview

Zubair Khalid

School of Science and Engineering
Lahore University of Management Sciences

https://www.zubairkhalid.org/ee514_2021.html

LUMS
A Not-for-Profit University

Zoom Policies

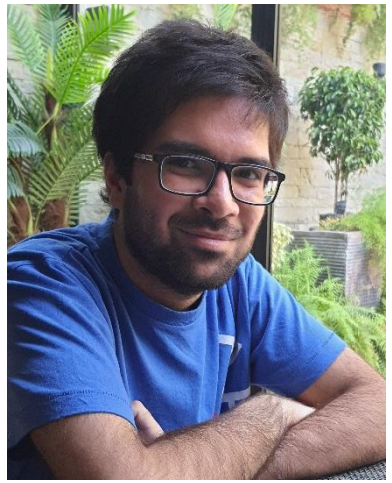
- Please use Zoom 'raise hand' feature if you have any questions.
- You will be asked to unmute yourself so that you can ask questions or provide your response.
- Please keep your microphone **mute at all other times.**
- We **do not** require you to turn-on your video.
- You can also post questions in the Chat. We may respond to the questions offline.

Zoom sessions will be recorded for offline viewing. Please let us know if you have any issues due to the privacy concerns.

About us!

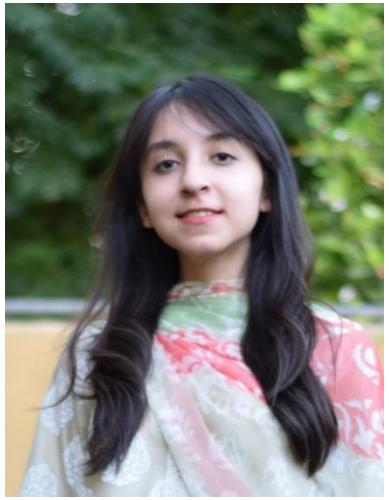
Zubair

Taimoor



Omer

Alishba



Rabeeya

Qasim

Mariyam

Haseeb

LUMS
A Not-for-Profit University

What is this course about?

Introductory course in Machine Learning (ML) – Fundamental topics in
- Supervised learning
- Unsupervised learning

Course Objectives:

- To provide a thorough introduction to ML methods
- To build mathematical foundations of ML and provide an appreciation for its applications
- To provide experience in the implementation and evaluation of ML algorithms
- To develop research interest in the theory and application of ML

Is this course a right choice for you?

Undergraduate students

- Interested in pursuing AI, Deep Learning and/or Machine Learning in their **grad school**
- Interesting in pursuing a **professional career** focused on the development of Machine Learning solutions

Graduate students

- Want to do fundamental research in the area of Machine Learning
- Wish to apply Machine Learning in their research work

Course Prerequisites

Undergraduate students

- Linear Algebra (MATH120)
- Probability (MATH230, DISC203, CS501)
- Programming (CS200, EE201)

Graduate students

- Encouraged to revise Linear Algebra and Probability concepts (on-the-fly)

We expect all the students to have good programming skills (in C/Python/MATLAB)

Note on Assignment 0!

Learning Interface

Delivery of Content:

Mode: Hybrid (Asynchronous/synchronous)

Lectures:

- Pre-recorded (uploaded on YouTube)
- Lectures for the week will be uploaded by Wednesday

Utilization of Lecture-slots:

- Tutorial in the Friday slot (@3:30 PM)
- Quiz (15-20 minutes)

Learning Interface

Communication:

Course Page: https://www.zubairkhalid.org/ee514_2021.html

Zoom: Same link for office hours, live sessions, meetings

Slack: Course-related questions or discussions. We will try to respond to the queries ASAP.

Office Hours: Posted on course page; distributed throughout the week

Email Policy:

Subject:

- 'ML-URGENT-Assignment Clarification'
- 'ML-NOT URGENT-Extend Assignment deadline'

Please **do not** flood my inbox to verify whether I have received your submission in LMS or the submission is late due to last-minute connectivity issues.

Grading Distribution

- **Programming Assignments and Homeworks: 30%**
  - 5 Programming Assignments
  - 3 Homeworks
- **Quizzes: 20%** (Almost every week)
- **Mid-exam or Mid-Viva: 15%**
- **Project: 10%**
- **Final Exam: 25%**

Course Polices

- **Homework Late Policy**
  - 10% per day for 3 days. No submission after 3 days (72 hours)
- **Missed Quiz Policy**
  - No make-up for quiz
- **Plagiarism will be strictly dealt with as per university policies (take it seriously).**
- **Zero Tolerance for Plagiarism and Cheating**
- Re-grading can be requested after grade reporting, within the following time limits:
  - HW and Assignments: 2 days
  - Final Exam: 3 days

Course Polices

Harassment Policy

Harassment of any kind is **unacceptable**, whether it be sexual harassment, online harassment, bullying, coercion, stalking, verbal or physical abuse of any kind. Harassment is a very broad term; it includes both direct and indirect behaviour, it may be physical or psychological in nature, it may be perpetrated online or offline, on campus and off campus. It may be one offense, or it may comprise of several incidents which together amount to sexual harassment. It may include overt requests for sexual favours but can also constitute verbal or written communication of a loaded nature. Further details of what may constitute harassment may be found in the LUMS Sexual Harassment Policy, which is available as part of the university code of conduct.

LUMS has a Sexual Harassment Policy and a Sexual Harassment Inquiry Committee (SHIC). Any member of the LUMS community can file a formal or informal complaint with the SHIC. If you are unsure about the process of filing a complaint, wish to discuss your options or have any questions, concerns, or complaints, please write to the Office of Accessibility and Inclusion (OAI, oai@lums.edu.pk) and SHIC (shic@lums.edu.pk) —both of them exist to help and support you and they will do their best to assist you in whatever way they can.

To file a complaint, please write to harassment@lums.edu.pk.

Course Polices

Help related to equity and Belonging at SSE

SSE's Council on Equity and Belonging is committed to devising ways to provide a safe, inclusive, and respectful learning, living, and working environment for its students, faculty, and staff.
For help related to any such issue, please feel free to write to any member of the school council for help or feedback.

Mental Health Support at LUMS

-

For matters relating to counselling, kindly email student.counselling@lums.edu.pk, or visit https://osa.lums.edu.pk/content/student-counselling-office for more information.

You are welcome to write to me or speak to me if you find that your mental health is impacting your ability to participate in the course. However, should you choose not to do so, please contact the Counselling Unit and speak to a counsellor or speak to the OSA team and ask them to write to me so that any necessary accommodations can be made.

Modules

1- ML Overview

Course Overview, notation
Supervised Learning Setup

Weeks: 1,2

Components:

- Programming Assignment 1: Intro to Python, Setting up Environment

Modules

2 - Classification

Classification
KNN
Evaluation Metrics, Curse of Dimensionality
Multi-class Classification

Weeks: 3,4

Components:

- Programming Assignment 2: KNN based (Using Images)
- Homework 1A

Modules

3 - Regression

Linear Regression
Gradient Descent
Multi-variate Regression
Polynomial Regression
Bias-Variance Trade-off, Regularization

Weeks: 4,5

Components:

- Programming Assignment 3: Regression
- Homework 1B

Modules

4 - Logistic Regression

Logistic Regression

Weeks: 6

Components:

- Programming Assignment 4: Logistic Regression

Modules

5 – Bayesian Framework

Bayes Theorem
Naive Bayes Classification

Weeks: 7,8

Components:

- Programming Assignment 5: Naïve Bayes Classifier (may be merged with Assignment 4)
- Homework 2

Modules

6 – Perceptron, SVM and Neural Network

Perceptron Algorithm
SVM
Neural Networks

Weeks: 9,10,11,12 **(Mid-Exam)**

Components:

- Programming Assignment 6: Neural Networks
- Homework 3

Modules

7 – Clustering

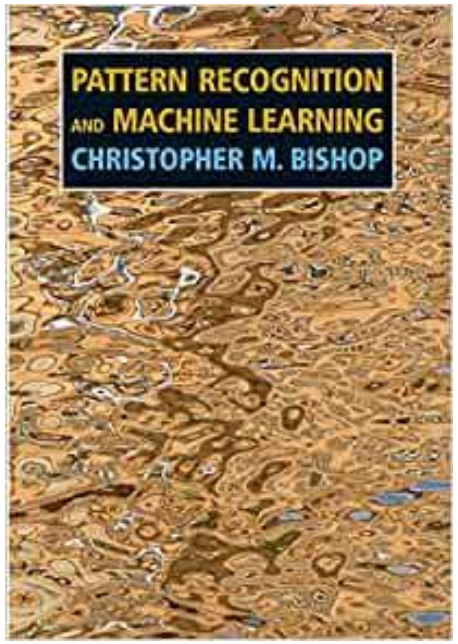
Unsupervised Learning Overview
Clustering (k-means)

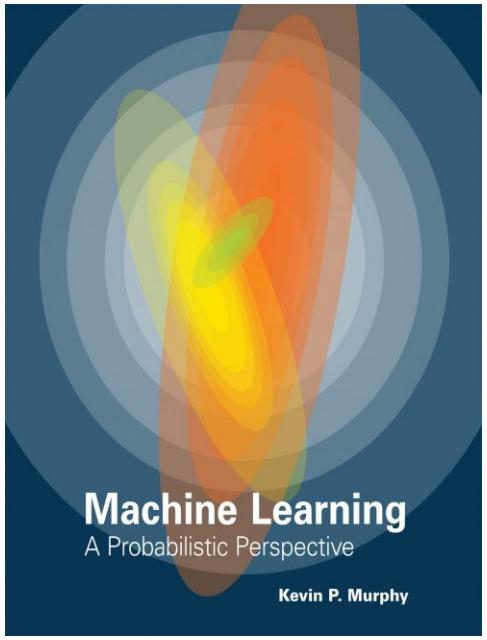
Weeks: 13,14

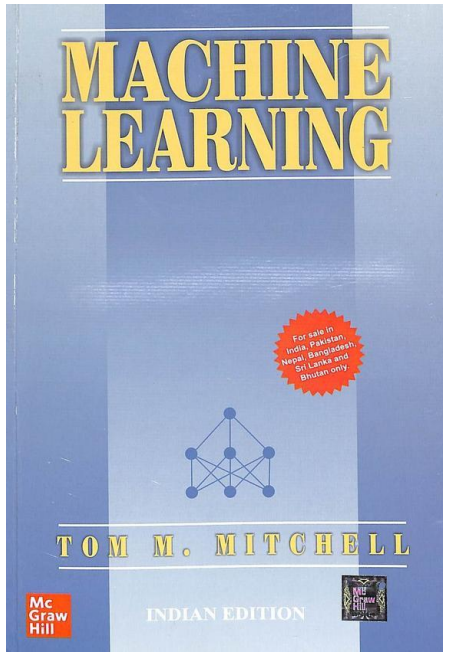
Components:

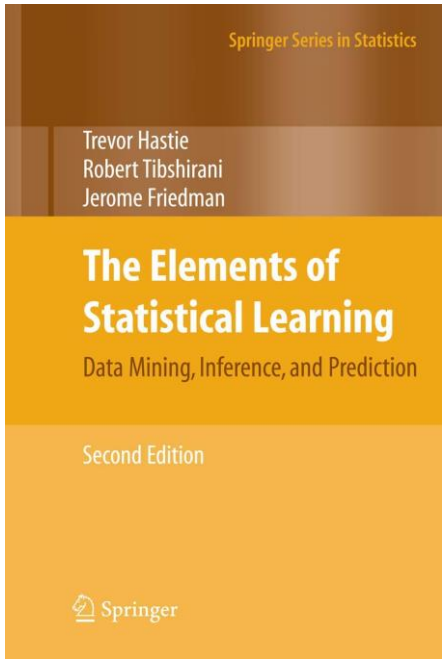
- Homework 3

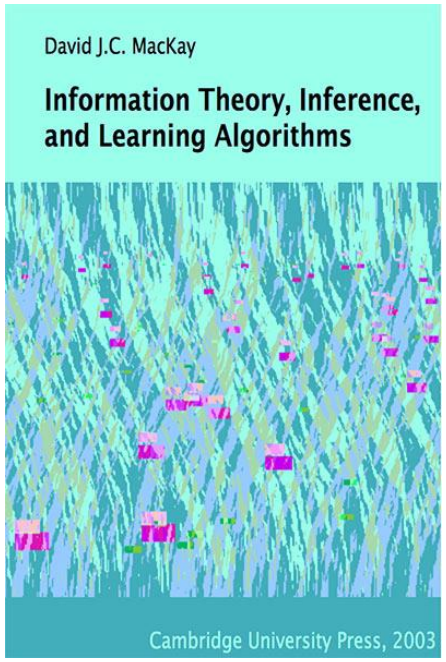
Suggested Reference Books











- (CB) Pattern Recognition and Machine Learning, Christopher M. Bishop
- (KM) Machine Learning: a Probabilistic Perspective, Kevin Murphy
- (TM) Machine Learning, Tom Mitchell
- (HTF) The Elements of Statistical Learning: Data mining, Inference, and Prediction, by Hastie, Tibshirani, Friedman
- (DM) Information Theory, Inference, and Learning Algorithms, David Mackay

"As to methods, there may be a million and then some, but principles are few. The man who grasps principles can successfully select his own methods."

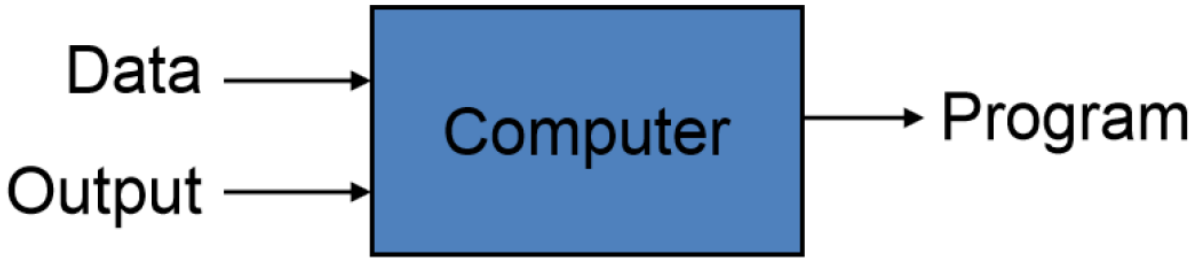
Ralph Waldo Emerson

Machine Learning Overview

What is Machine Learning?

- Automating the process of automation
- Getting computers to program themselves

Traditional Programming

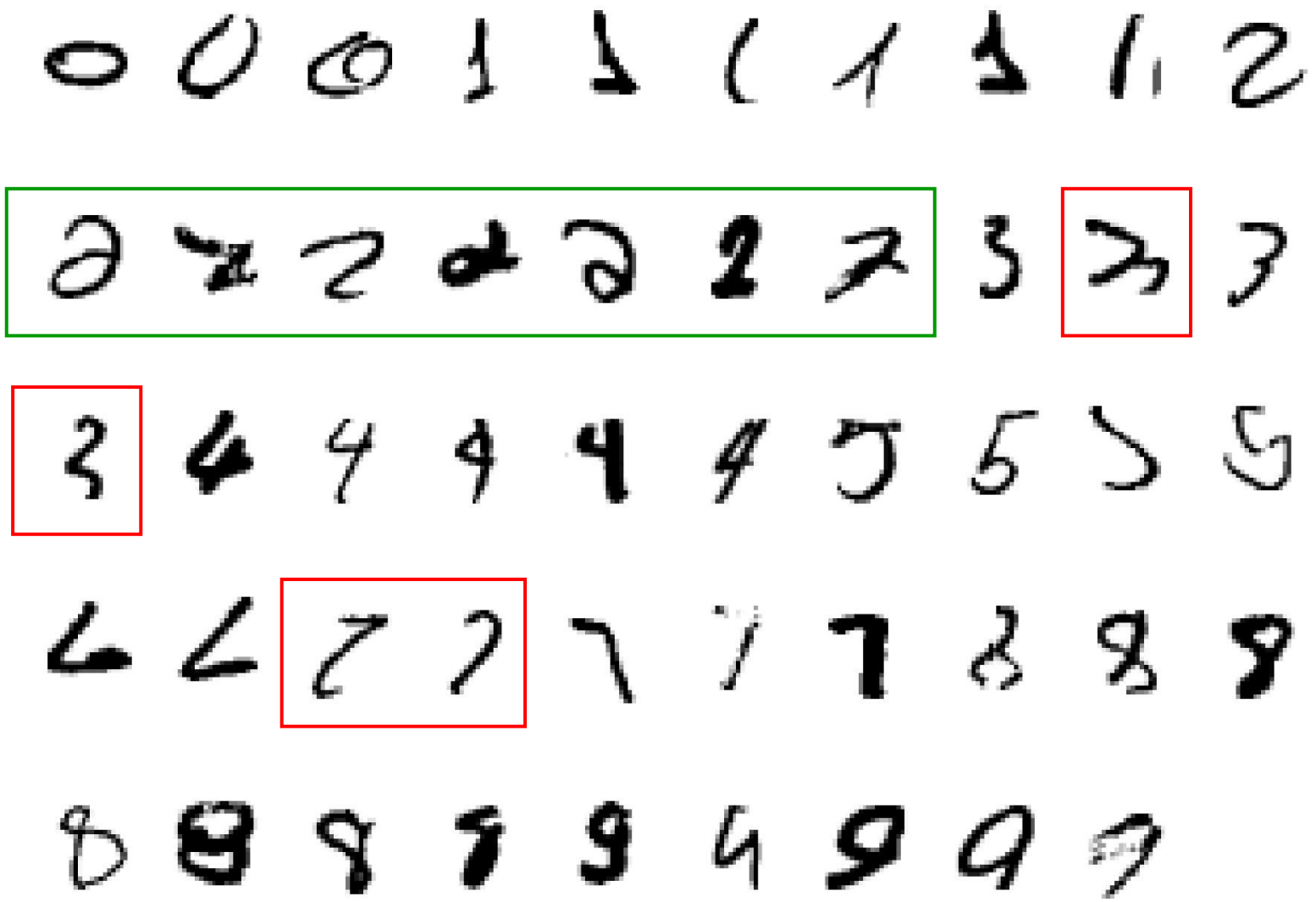


Machine Learning

Given examples (training data), make a machine learn system behavior or discover patterns

Machine Learning Overview

Classical Example: Recognize hand-written 2!



Machine Learning: Overview

Example Applications

- Medical Diagnosis
- Autonomous Driving
- Information extraction
- Computer/Machine Vision
- Finance
- Web Search
- Robotics
- Social networks
- Production Industry
- Logistics
- Waste Management
- [Your research/favorite area]

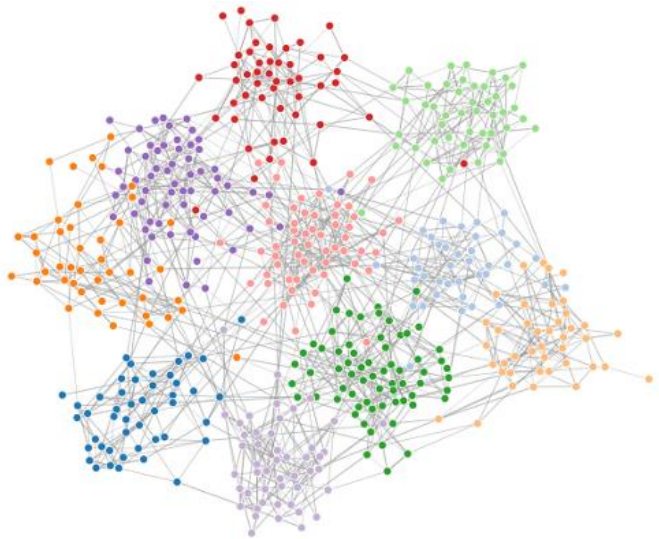
Face Recognition: Demo

Machine Learning: Overview

What is Machine Learning?

- *Study, Design and Analysis of algorithms that improve their **performance** at some task with **experience***

Data



Given to us

Algorithm

We need to design it

Model

$$f(\mathbf{x})$$

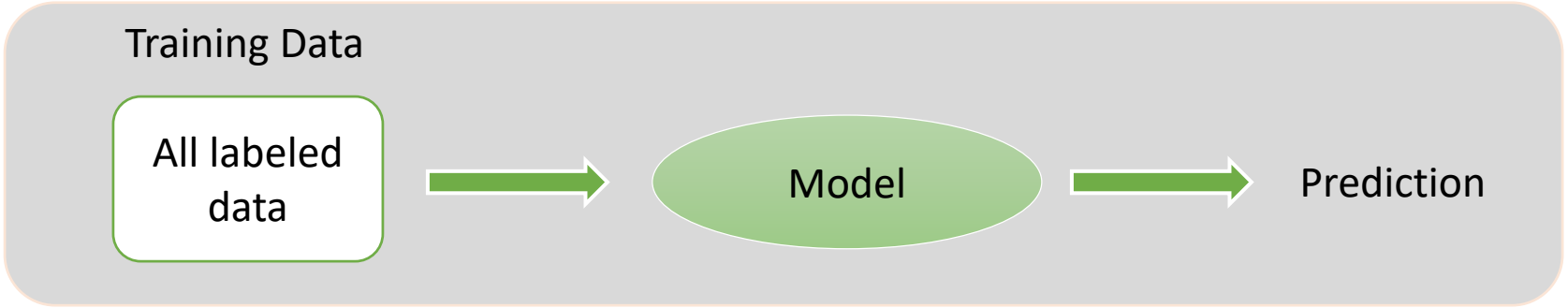
Final output which enables us to make predictions

Machine Learning: Overview

Nature of ML Problems

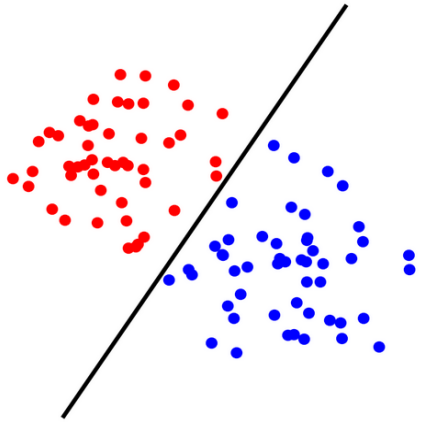
1. Supervised Learning

The learning algorithm would receive a set of inputs along with the corresponding correct outputs to train a model



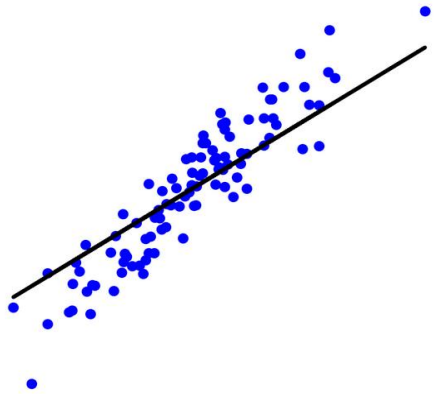
Classification: Discrete Prediction

Given a data sample, predict its class



Regression: Quantitative Prediction on a continuous scale

Given a data sample, predict a numerical value

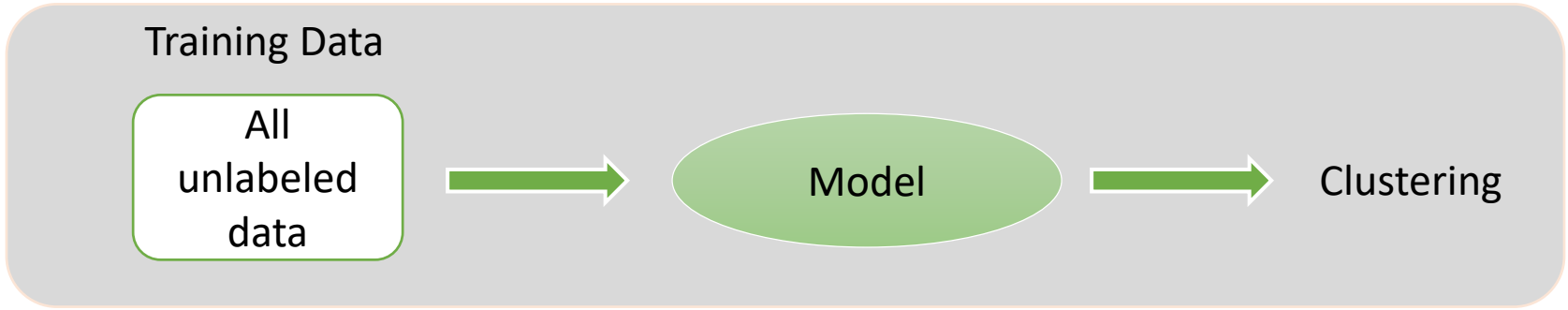


Machine Learning: Overview

Nature of ML Problems

2. Unsupervised Learning

The learning algorithm would receive unlabeled raw data to train a model and to find patterns in the data

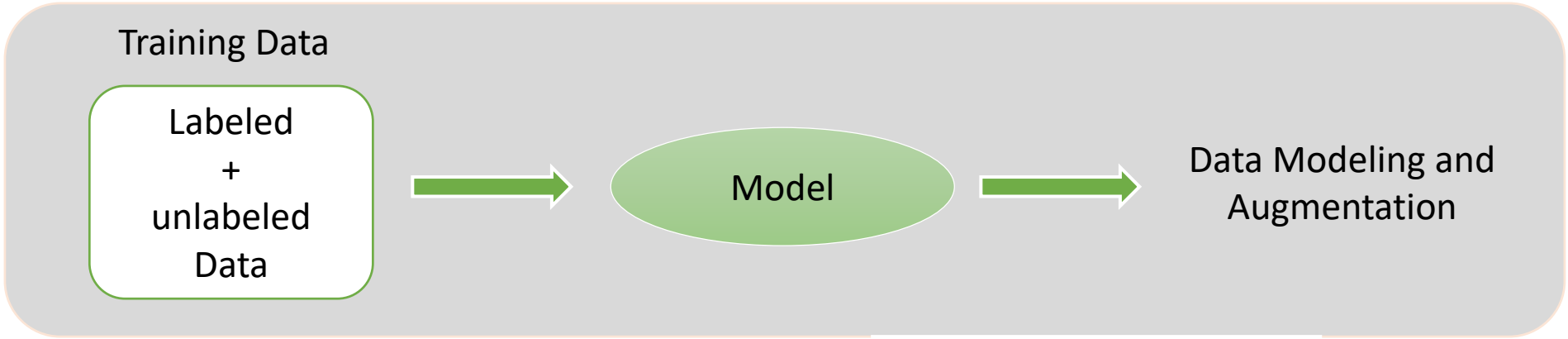


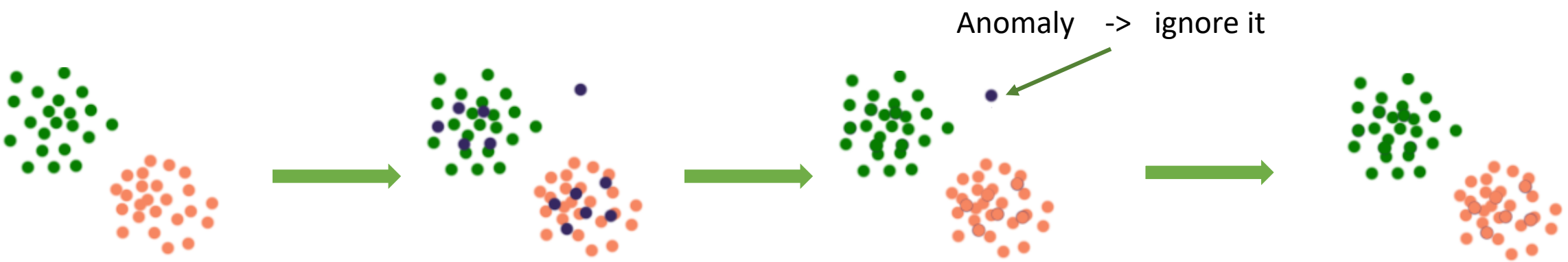
Machine Learning: Overview

Nature of ML Problems

3. Semi-supervised Learning

- *The learning algorithm receives labeled and unlabeled raw data to train a model*
- *Main objective is to efficiently accommodate the unlabeled data*





Use labeled data to build a model

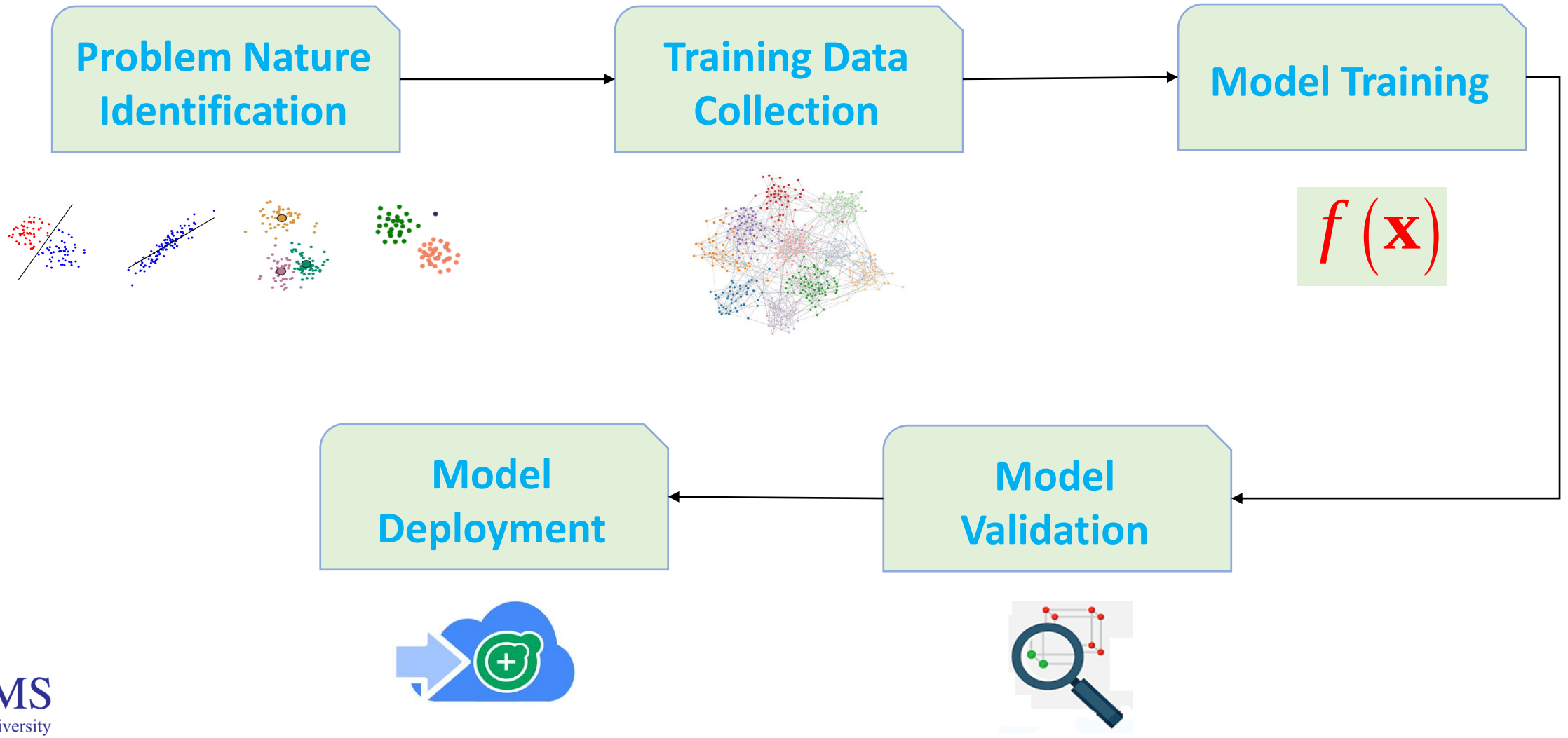
Place unlabeled data with model

Use the model to label the unlabeled data

Fit the model again with the combined data

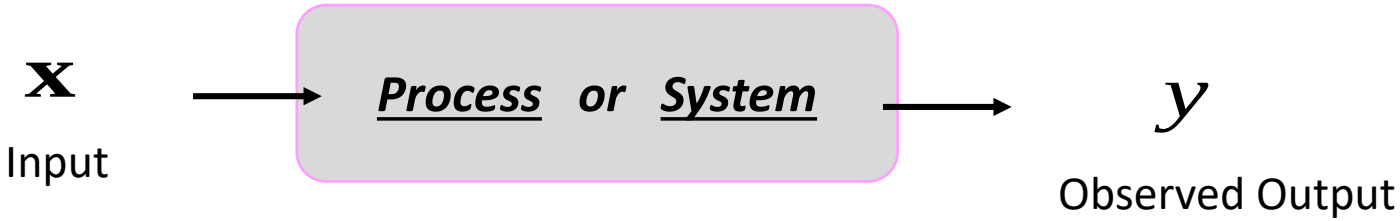
Machine Learning: Overview

Typical Flow



Machine Learning: Overview

Training Data Collection

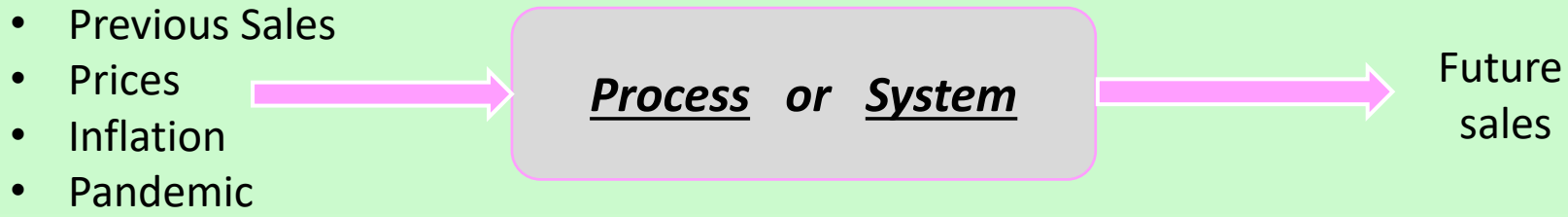


PROCESS or ***SYSTEM*** : Underlying physical or logical phenomenon which maps our input data to our observed output

Collect the training data by observing our unknown **PROCESS** or **SYSTEM**

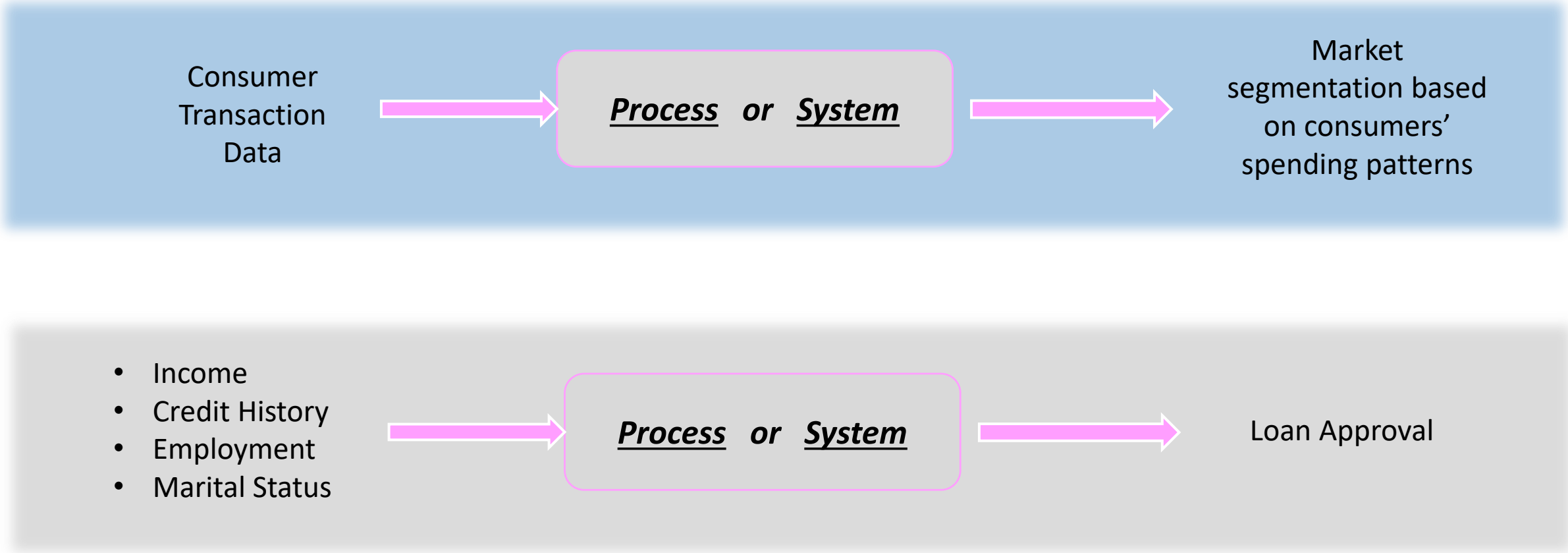
Machine Learning: Overview

Example Systems



Machine Learning: Overview

Example Systems

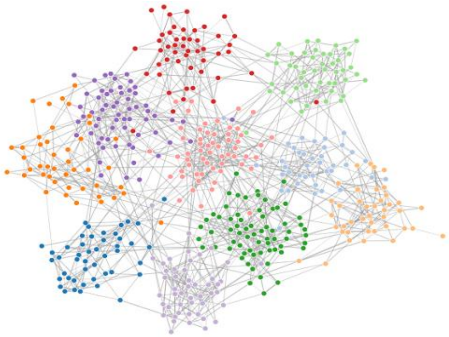


Machine Learning: Overview

Data Types

Tabulated Data

- *Each column is a feature and adds one dimension to the data*
- *Number of columns define total number of features and hence data dimensionality*
- *Ideally, these columns should contain different information -> uncorrelated*

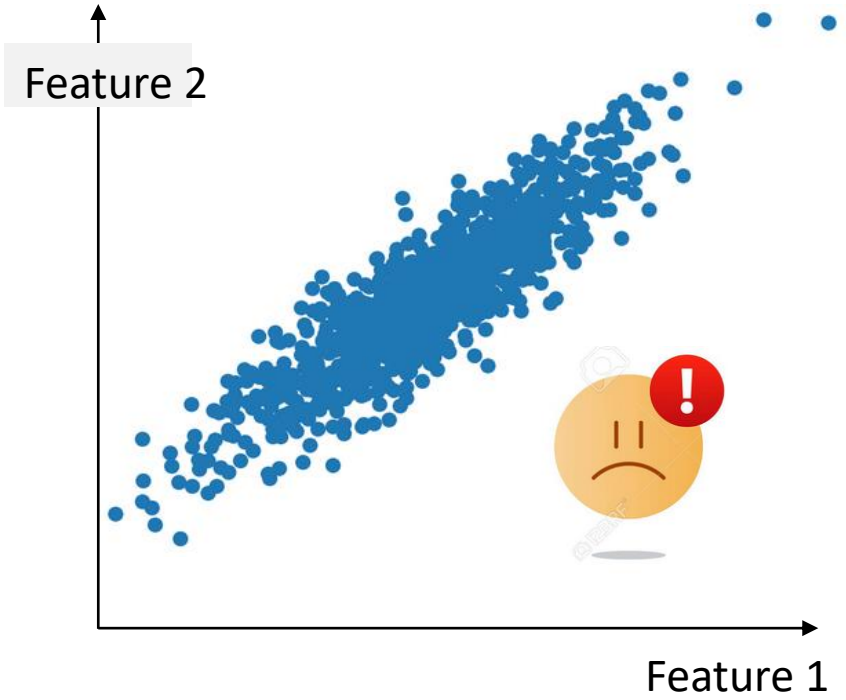


rows 32561 columns 15

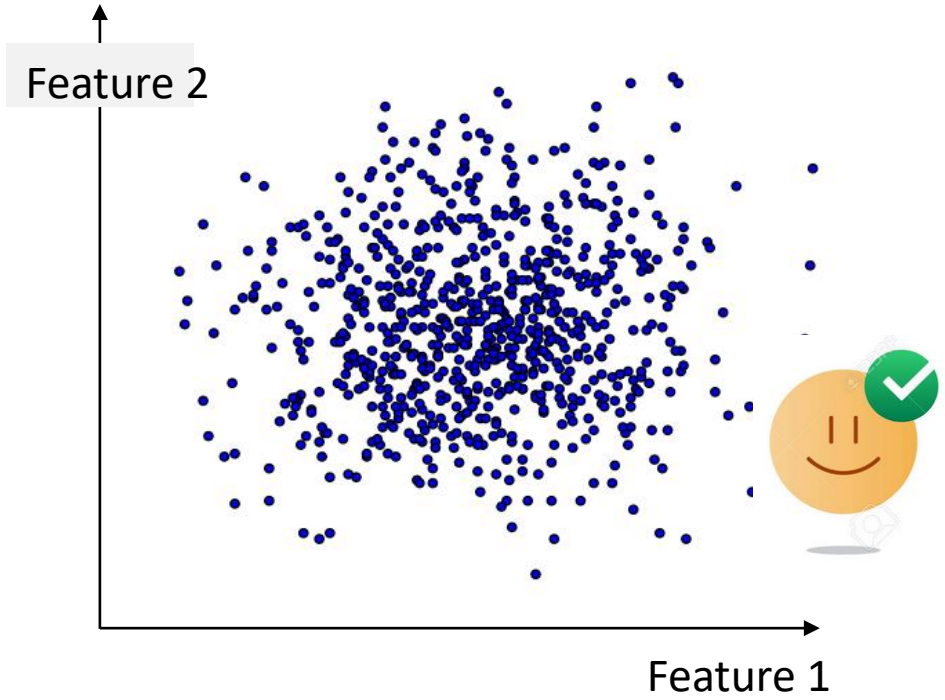
	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
view as															
	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=50K
	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=50K
	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=50K
	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=50K
	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	<=50K
	37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	White	Female	0	0	40	United-States	<=50K
	49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family	Black	Female	0	0	16	Jamaica	<=50K
	52	Self-emp-not-inc	209642	HS-grad	9	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	45	United-States	>50K

Machine Learning: Overview

Data Types



Features are correlated



Features are uncorrelated

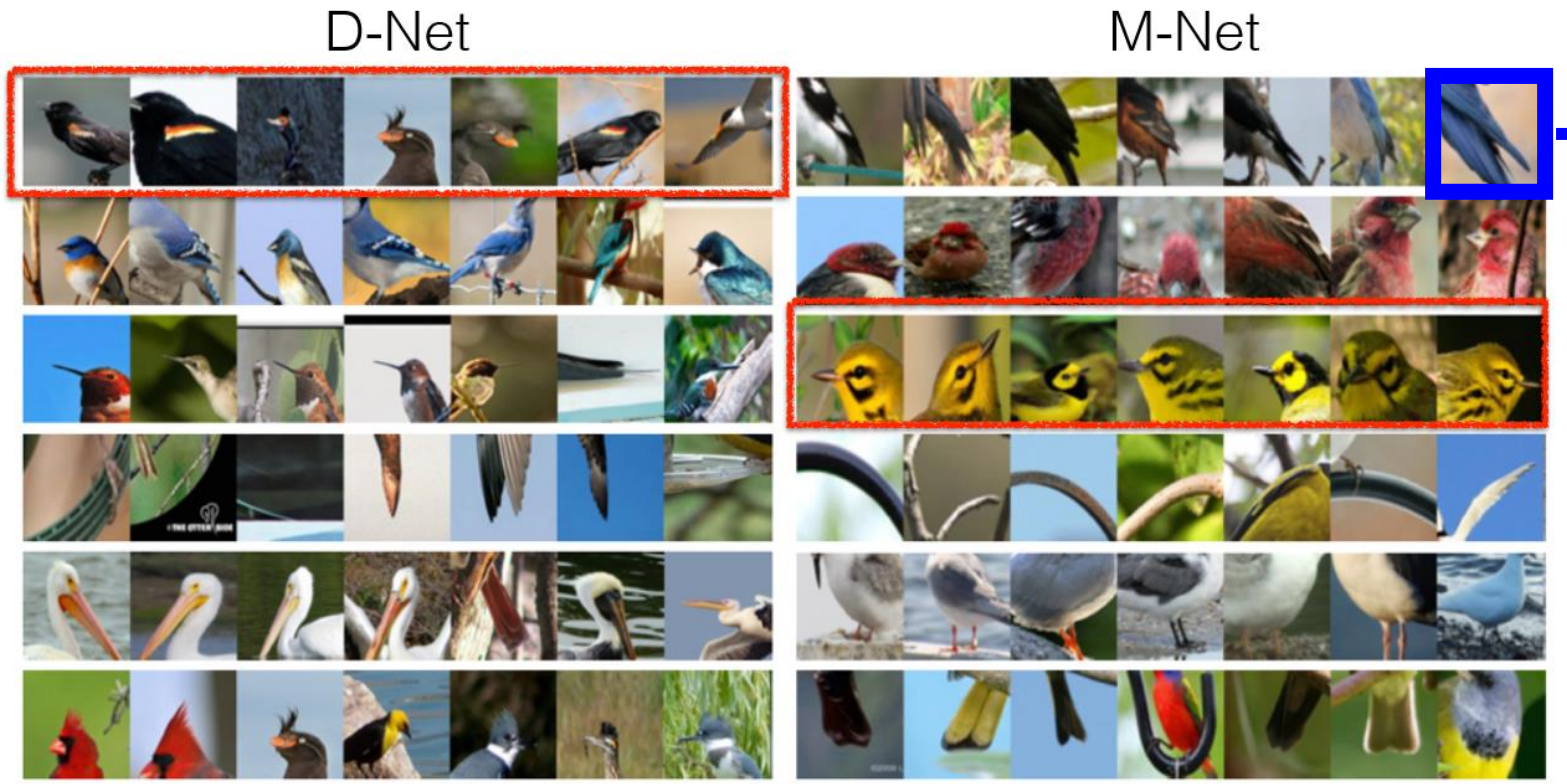
Dimensionality Reduction

Find all those dimensions (features) which carry very less information and discard them

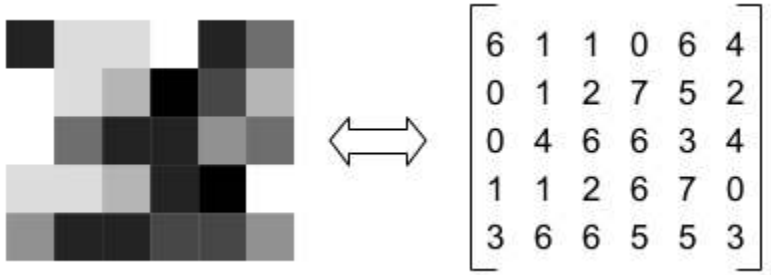
Machine Learning: Overview

Data Types

Images



- Each image (greyscale) is a 2D data which can be represented as a matrix



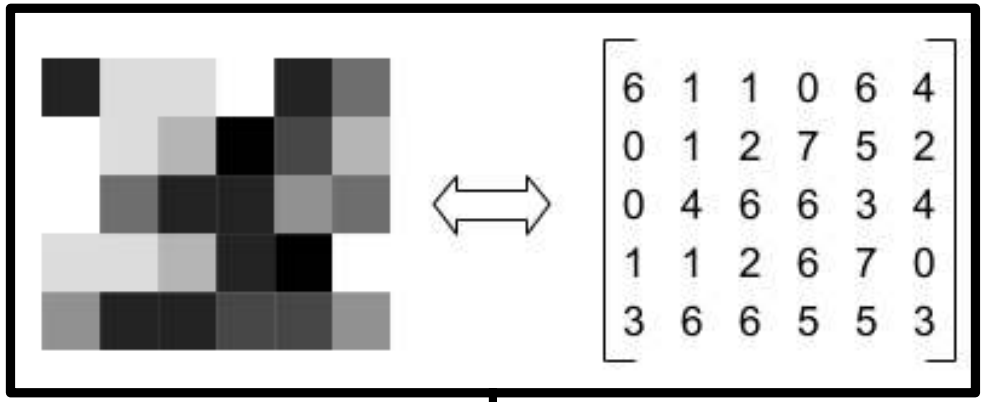
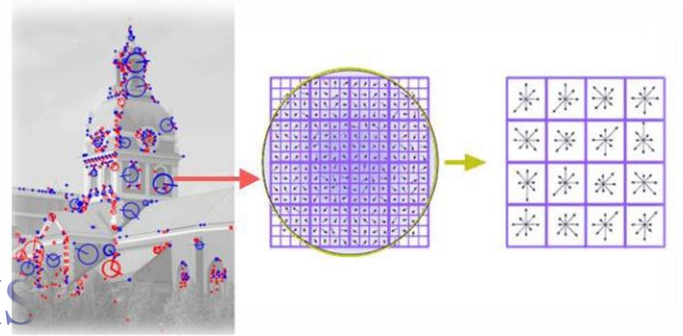


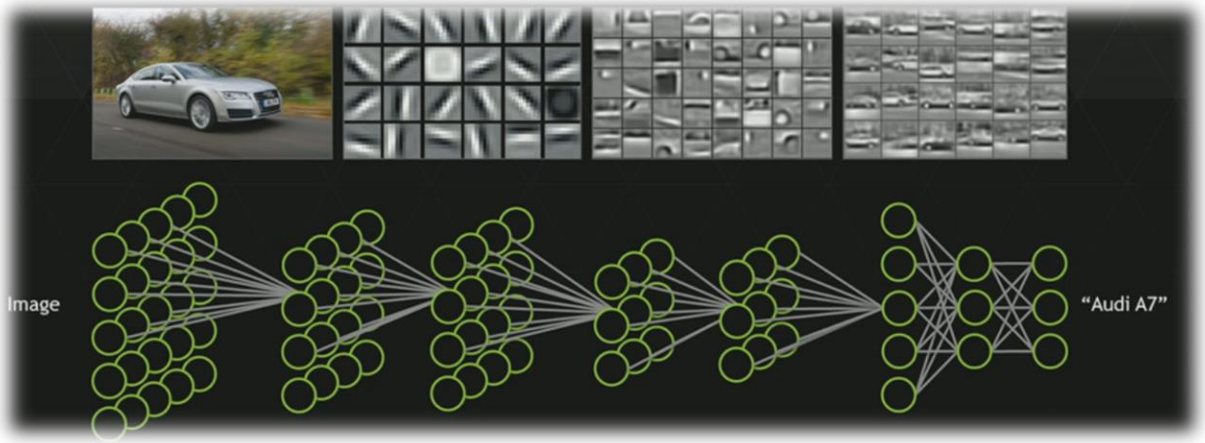
Image Descriptors: Manually extract features

- Speeded-Up Robust Features (SURF)
- Binary Robust Independent Elementary Features (BRIEF)
- Oriented FAST and Rotated BRIEF (ORB)
- Scale-Invariant Feature Transform (SIFT)
- Histogram of Oriented Gradients (HOG)
- HAAR features
- etc.



Deep Learning: Let machine learning extract most important features

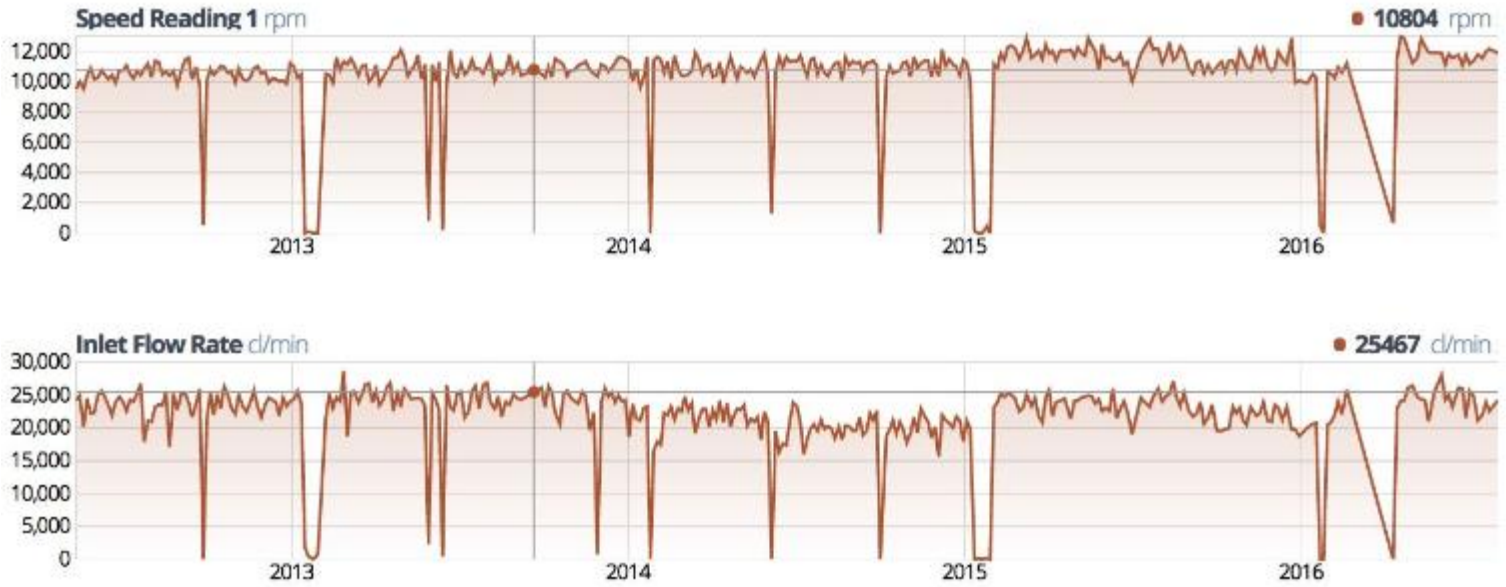
Neural Networks such as Convolutional Neural Networks (CNN) take raw images as input and first process it inside multiple layers to get best features called intrinsic dimension of the data



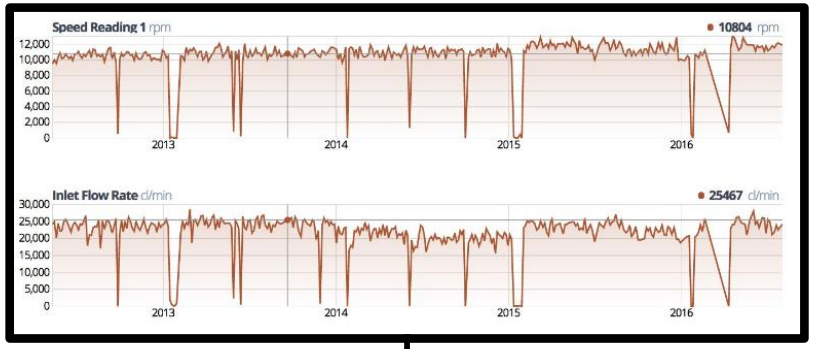
Machine Learning: Overview

Data Types

Time Series

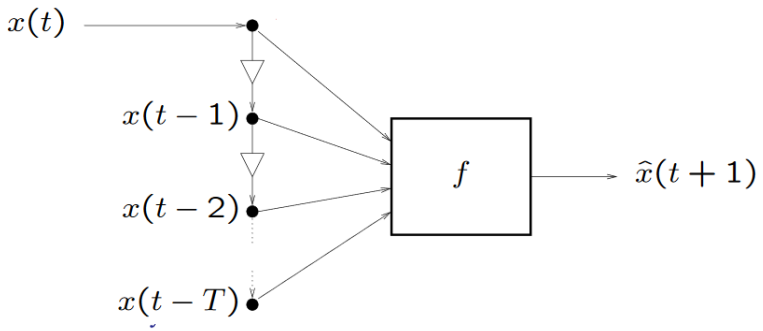


- Much difficult to deal with
  - Enormous Size
  - Cumbersome Labeling
  - Not so standard feature extraction strategies as in case of images. (Direct modeling of time series is preferred)

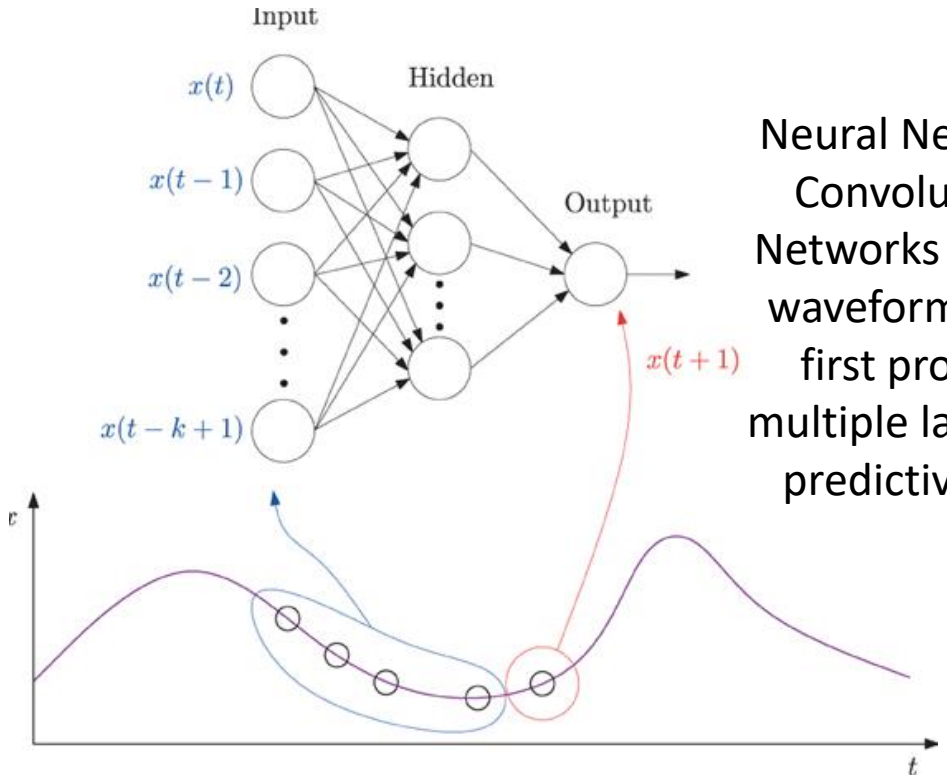


Time series Models: Manually model it !

- Exponential Smoothing (ETS)
- Autoregressive Integrated Moving Average (ARIMA) models
- Linear regression
- Generalized Autoregressive Conditional Heteroskedasticity (GARCH)
- Bayesian Models
- Vector auto-regression (VAR) models
- etc.



Machine Learning: Let machine learning handle the task !

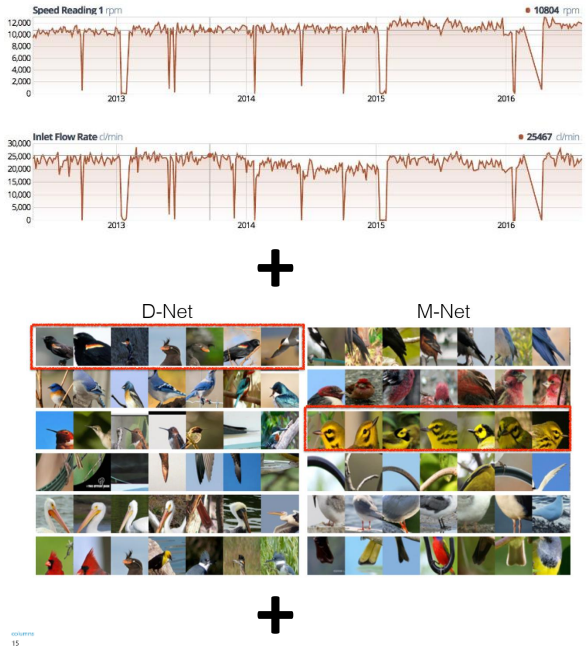


Neural Networks such as Convolutional Neural Networks (CNN) take raw waveforms as input and first process it inside multiple layers to get best predictive model for it

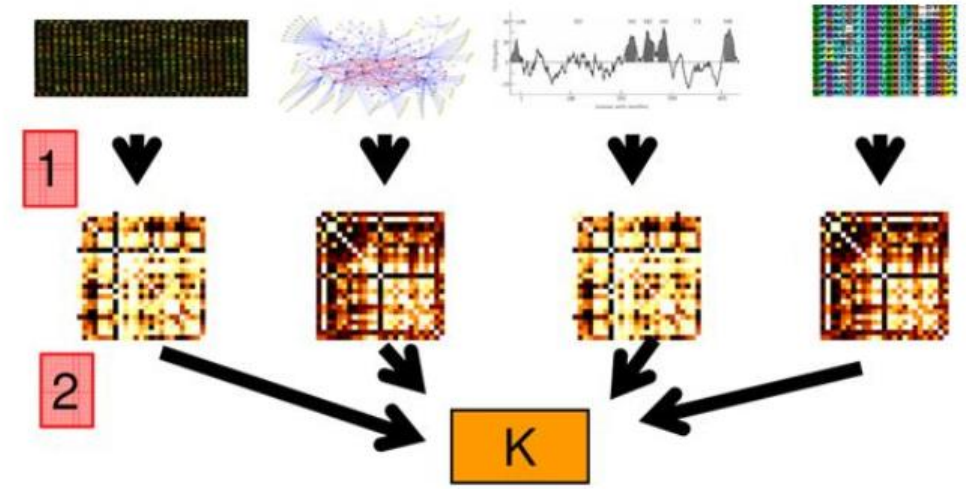
Machine Learning: Overview

Data Types

Heterogeneous data



Multimodal Machine Learning !



LUMS
A Not-for-Profit University

About the Instructor

- Assistant Professor, LUMS since 2015
- Post-doctorate – 2013-2015, Australian National University (ANU)
- PhD, Australian National University (ANU) - 2013

Affiliations:
- Signal, Image and Video Processing Lab, LUMS
- Applied Signal Processing Group, ANU
- Smart Data, Systems and Applications Lab (www.sdsa.lums.edu.pk)

Collaborations: Princeton, UCL, University of Edinburgh, EPFL, ANU

PhD Students: 7 (4 graduated)

Publications: More than 70 (21 Transactions/Journals, 50 Conference proceedings)

Service: Senior Member IEEE and Associate Editor, IEEE Signal Processing Letters

Feedback: Questions or Comments?

Email: zubair.khalid@lums.edu.pk